A NOTE OF CAUTION ON COVARIANCE-EQUIVALENT MODELS IN INFORMATION SYSTEMS

Completed Research Paper

Joerg Evermann
Memorial University of Newfoundland
St. John's, Canada
jevermann@mun.ca

Mary Tate
Victoria University of Wellington
Wellington, New Zealand
mary.tate@vuw.ac.nz

Abstract

Covariance-based structural equation modeling is a popular statistical technique in information systems research, providing a stringent test of model fit and allowing researchers to test multiple hypotheses in the same model. Structural regressions in such models are often assumed to represent the causal nature of the underlying reality as expressed by theory. The validity of conclusions drawn from covariance-based analysis is, however, challenged when models can be constructed that fit the observed covariances equally well as the tested model, but which have a different structure, expressing different underlying causal relationships. This research shows that a large proportion of studies in IS exhibit this issue. The dangers posed by covariance-equivalent models are highlighted using an example in the published literature, and recommendations are provided to IS researchers to address the problem.

Keywords: Covariance analysis, equivalent models, validity

Introduction

Covariance-based analysis of structure equation models is a frequently used statistical technique in information systems (IS) research. It is used to test latent variable models that represent causal relationships among theoretical concepts. However, multiple, semantically different models may be equivalent with respect to their fit with the observed covariance data. In those cases, it is impossible to decide between such equivalent models based on the covariance information. Thus, the internal validity of the study, i.e. the ability to rule out alternative hypotheses (Straub, Boudreau, and Gefen, 2004), is threatened.

A survey of four highly-ranked IS journals (MISQ, ISR, JMIS, JAIS) between 2004 and 2010 (inclusive) identified 61 studies using covariance-based analysis for testing the main theoretical model. This does not include studies that use covariance analysis for confirmatory factor analysis (CFA) models but move to regression models or partial least squares analysis for hypothesis testing. Despite Chin's (1998) warning to IS researchers, none of the 61 studies discuss the possible threat to the validity of the conclusions drawn from the statistical results that is posed by covariance-equivalent models. When examining these studies closely, we found that more than one in three (22 of 61) of them exhibit problems with covariance-equivalent models. This motivates us to raise the issue.

The purpose of this paper is not to critique specific past studies (though we use them to illustrate the dangers that equivalent models can pose), but to bring the issue to the attention of IS researchers. We provide recommendations so that the validity of substantive conclusions from covariance-based studies can be strengthened in the future.

The remainder of this paper is structured as follows. The next section discusses covariance analysis and model equivalence. This is followed by an illustration of the threats to the validity of theoretical conclusions that is posed by model equivalence. For this, we examine one published study in detail and briefly point out issues with other studies we have identified. We conclude with recommendations for researchers to avoid this issue in the future.

Covariance Analysis and Equivalent Models

Covariance analysis is based on the idea that structural equation models, which may include latent and observed variables linked by linear regression equations, imply a covariance matrix of the observed variables. This implied covariance matrix is a function of the free model parameters (regression coefficients, exogenous variances and covariances, including error variances and covariances) and model constraints. Typical constraints are “missing” paths, i.e. regression coefficients fixed to zero, equality constraints, etc. Covariance analysis estimates free parameters by minimizing the difference between the model-implied and the observed covariance matrix in the free parameters, subject to the model constraints.

The problem of equivalent models arises because different structural equation models can imply the same covariance matrix of the observed variables. While this is only a problem when these models have different substantive, theoretical interpretations (Markus, 2002), the following section shows that this is the case with many of the problematic models in the IS literature.

The possibility of equivalent models has been recognized in the psychometrics literature, where procedures for generating equivalent models have been proposed based on replacement of model elements in saturated (i.e. fully-connected) model blocks (Stelzl, 1986; Lee and Hershberger, 1990). A different approach to constructing equivalent models, based on feedback loops, has been proposed by Hayduk (1996). More recently, a method for generating an infinite number of equivalent models for any given model has been proposed, based on decomposing exogenous variables into orthogonal components (Raykov and Marcoulides, 2001). However, there is debate whether this last procedure yields models that are semantically equivalent, i.e. models that have the same substantive, theoretical interpretation. If that were the case, model equivalence would not be problematic (Markus, 2002).

We now illustrate some simple equivalent models, based on the replacement rule by Lee & Hershberger (1990). We do not aim to be comprehensive but refer the reader to the very accessible treatment in the recent textbook by Mulaik (2009).

The main idea behind the replacement rule, which subsumes the earlier method by Stelzl (1986), is to identify saturated blocks in the model. Saturated blocks are those in which each variable is directly connected to every other variable, either by a regression relationship, by a covariance, or by a correlation of disturbance/error terms. If all variables in such a saturated block have the same set of antecedent variables, any or all of the direct connections within the block may be replaced by any other type of direct connection, including a reverse regression relationship (Figure 1). An example saturated block is shown in Figure 2. Variables X and Y are directly connected with each other and have the same antecedent variables A and B. Thus, the regression relationship between the two may be replaced by any other direct relationship, as shown in Figure 1. The models in Figures 3 (reversing the regression between X and Y) and 4 (replacing the regression with correlated disturbance terms) are derived by this rule. This rule can be applied iteratively to the resulting covariance-equivalent model, possibly yielding further equivalent models.

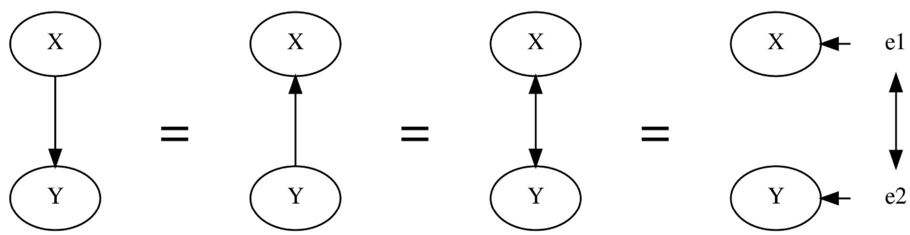

Figure 1: Replacement rule (from Mulaik, 2009)

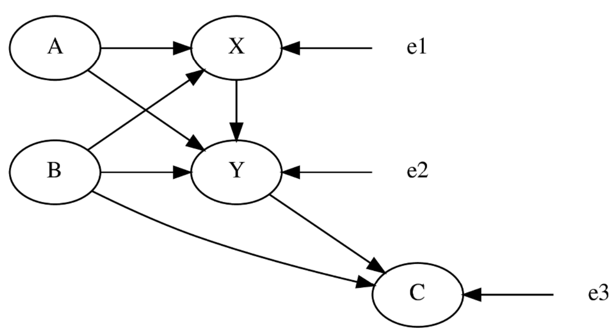

Figure 2: Example model with saturated block with common antecedents (from Mulaik, 2009)

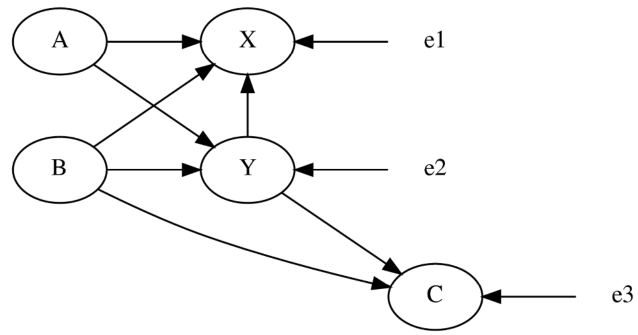

Figure 3: Equivalent model to Figure 2 (reversal of regression between X and Y)

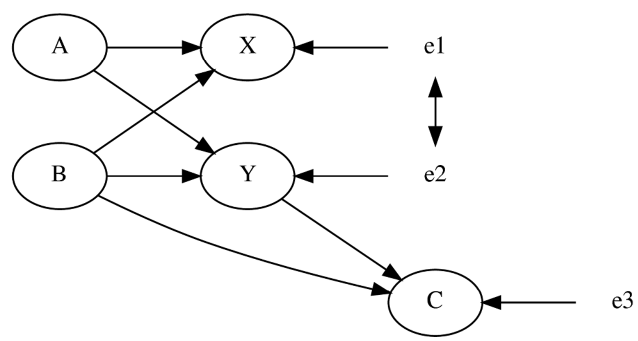

Figure 4: Equivalent model to Figure 2 (correlated disturbances of X and Y)

As a corollary, if the block of common antecedent variables is exogenous and also saturated, the combined model block is saturated with the same set of antecedents (the empty set). For example, if Figure 2 contained a free covariance between A and B (which is a frequent assumption for covariance models), the entire block consisting of A, B, X, and Y is saturated and any or all of its connections may be replaced according to Figure 1.

Figure 5 shows some of the 12 equivalent models for a simple three variable model. The Technology Acceptance Model, one of the most widely used models in IS, is of this type. As should be clear from the requirement of saturated blocks for the replacement rule, less constrained models, i.e. those with large numbers of free parameters or paths, are more susceptible to the problem of equivalent models than more constrained models, i.e. those with small numbers of free parameters or paths.

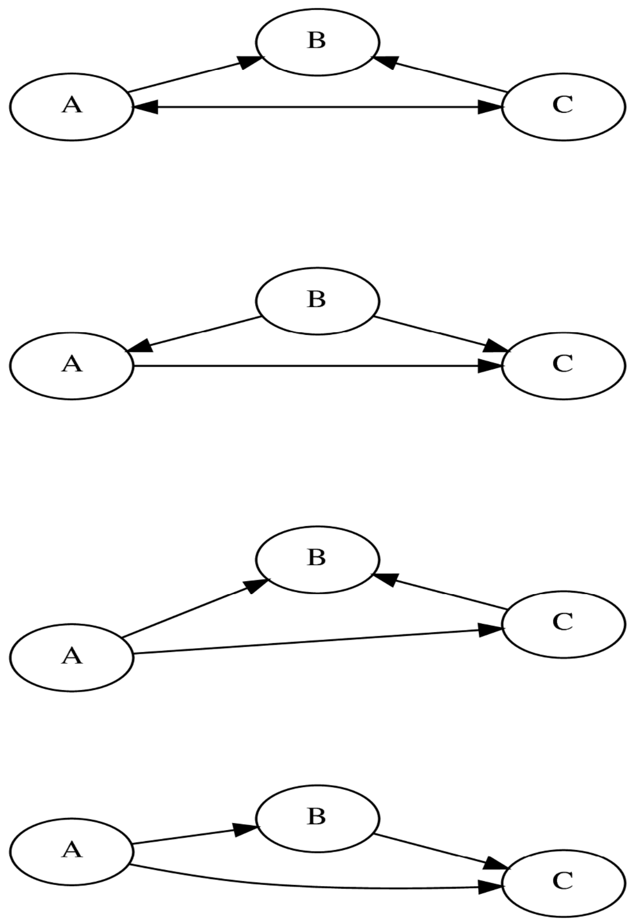

Figure 5: Four of 12 equivalent models for three variables

Equivalent Models in Information Systems

Detailed Example

Of the 22 studies that we identified in the four journals (MISQ, ISR, JMIS, JAIS) as exhibiting problems with equivalent models, three provide covariance matrices (Pavlou and Gefen, 2004; Li, Chau, and Lou, 2005; Chin, Johnson, and Schwarz, 2008) so that we are able to illustrate the threats to theory testing posed by equivalent models. The TAM model considered by Chin et al. (2008) is relatively simple, and we were unable to reconstruct the model tested by Pavlou and Gefen (2004) based on the published information. Hence, we focus on the model by Li et al. (2005) shown in Figure 6.

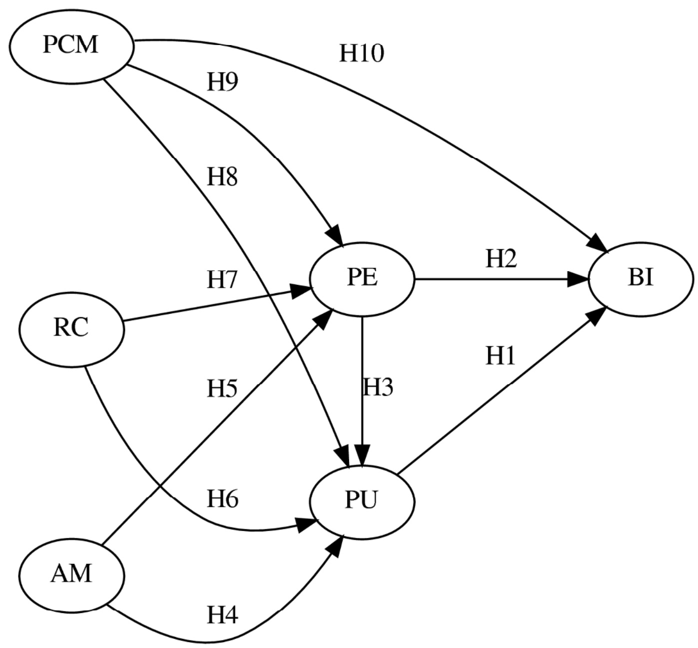

Figure 6: Research Model and Hypotheses from Li, Chau, and Lou (2005)

The model block in question is the regression path from perceived usefulness (PU) to perceived enjoyment (PE), representing Hypothesis 3. Both PU and PE have the same set of antecedent variables. Li et al. (2005) posit "that the fun, pleasure, and enjoyment derived from using IM with friends will positively affect an individual's perception of the usefulness of IM" (pg. 110), citing a previous study by Venkatesh (2000). There, PE is argued to influence perceived ease of use (not included in the Li et al. model), which in turn is argued to influence PU: "We expect that with increasing experience, system use may become more routinized, less challenging, and less discovery-oriented. In such cases, the lack of enjoyment may cause system use to be perceived to be more effortful" (Venkatesh, 2000, pg. 351). Thus, the proposed causal chain in support of H3 is that systems that are perceived to be enjoyable to use will also be perceived as easy to use and systems that are perceived as easy to use will be perceived as being useful.

Li et al. (2005) report that H3 is supported by a statistically significant unstandardized parameter estimate of 0.53. Additionally, they report that their H1, H2, H6, H7, H8, H9, and H10 are supported (Table 1). We replicated the results of Li et al. (2005) using the provided covariance matrix, but were unable to arrive at identical estimates. Nonetheless, our results support the same hypotheses at the same significance levels as reported by Li et al. (2005)[1]. Instead of the Li et al. (2005) numbers, columns two and three in Table 1 report our own numbers.

1 Differences are possibly be due to rounding of values in the reported covariance matrix, treatment of missing values, estimator used in the CB-SEM model, software used for the estimation, and other characteristics of the sample or the estimation not reported by Li et al. (2005). Our model had the same degrees of freedom (176) as reported by Li et al. (2005) supporting the correctness of our model. We used the reported sample size of 273.

Noting that both PU and PE have the same set of antecedent variables, a covariance-equivalent model can be specified by reversing the path between PU and PE. This represents the alternative causal chain that systems are perceived to be enjoyable when they are perceived to be useful. This appears also plausible[2] as in this case, the enjoyment of working with the system is derived from the accomplishment of the task that the system is intended for: When a system allows me to do my task, I enjoy using it more than if it didn't.

The equivalent model yields, by definition, an identical $\chi 2$ model fit test statistic. The parameter estimates and their significance are shown in Table 1 and compared to those of the original model. Note that H3 in the equivalent model is the hypothesis that perceived usefulness positively influences perceived enjoyment.

Comparing the two sets of parameter estimates, we notice first that both models support a significant path between PE and PU, though the theoretical conclusions are very different. Second, we notice that while the original model supports H7 but not H6, the equivalent model supports H6 but not H7. These changes clearly have important theoretical implications and affect the conclusions that Li et al. (2005) draw from their study. Finally, the parameter estimate for H9 was approximately twice the parameter estimate for H8 in the original model. In the equivalent model, the parameter estimate for H8 is approximately twice the parameter estimate for H9. While Li et al. (2005) do not discuss the relative strength of the relationship in theoretical terms, these changes also have theoretical implications.

Hypothesis	Path Coefficient in Li et al. (2005)	Significance in Li et al. (2005)	Path Coefficient in equivalent model	Significance in equivalent model
H1	.399	**	.399	**
H2	.401	**	.401	**
H3	.547	**	.457	**
H4	-.128		.042	
H5	.311	**	.292	**
H6	.123		.201	**
H7	.143	*	.051	
H8	.225	**	.447	**
H9	.407	**	.203	**
H10	.194	**	.194	**

Table 1: Parameter estimates in original and equivalent model of Li, Chau, and Lou (2005) (= Significant at p<0.001; * = Significant at p<0.05)**

Further, because there are free covariances between the exogenous latent variables in the model (attachment motivation AM, relationship commitment RC, perceived critical mass PCM), all variables but BI (behavioural intention) are part of a block in which any or all links can be replaced according to Figure 1 to yield equivalent models. With four connections between variables for each of the five variables (AM, RC, PCM, PE, PU) and four ways of modeling the connection (Figure 1), there exist 4^(4+5) possible

2 Without critiquing in detail every study from Venkatesh (2000) backwards, support for H3 by Venkatesh was ambiguous at best as the model was estimated using PLS, a technique known to have no test of overall model fit. Further, the Venkatesh model incorporates the Technology Acceptance Model in which the influence of Perceived Ease of Use on Usefulness has been questioned. In fact, the model in Chin et al. (2008), noted above for exhibiting equivalent models, does not hypothesize such a relationship.

equivalent models. Not all of them are theoretically sensible, helped in this case by the fact that attachment motivation (AM) and relationship commitment (RC) represent innate personality characteristics, whereas PE and PU represent specific beliefs about a system. It appears more plausible to suggest that innate characteristics affect situational beliefs than vice versa.

Our assumption of reversed causality between PE and PU in the model by Li et al. (2005) is not only plausible to us. The model of online impulse buying by Parboteeah, Valacich, and Wells (2009) is directly related to our detailed example, as their study concludes that perceived usefulness affects perceived enjoyment, in contrast to the conclusions by Li et al. (2005). This is a clear demonstration of the uncertain and ambiguous causality that can arise as a result of equivalent models. The conclusions by Parboteeah et al. (2009) are also uncertain as the relationship between PU and PE in their model is part of a saturated block that also includes the constructs Task-relevant cues and Model-relevant cues. This casts doubt not only on their conclusions about their H3 (PU and PE), but also on their hypotheses H1A-H1C and H2A-H2C. These are hypotheses about the relative strength of effects of the antecedent variables. As we have shown in the detailed example above, the parameter estimates, and thus their relative strength, can change in covariance-equivalent models.

To escape the problem of covariance equivalent models, the researcher might simply omit one of the problematic paths and fix it to a zero value (or perhaps some other value based on prior literature). For example, the model in Figure 6 where H4 is removed is no longer covariance equivalent to the model in which then H3 is reversed. It turns out that the latter model, where PE is regressed on PU fits better than the model that regresses PU on PE. In this case, we know post-hoc that the path is non-significant, but of course in a real research setting, the path omission needs to be specified before "peeking" at the data like this.

Further Examples

As we do not have covariance matrices available to illustrate the specific effects of equivalent models on parameter estimates, we briefly note issues with some of the remaining studies we have identified.

Pavlou and Gefen (2004) present a study on the effect of institutional trust mechanisms on trust, risk, and transaction intentions in electronic commerce. In their model, the constructs perceived risk and transaction intentions have the same antecedents, trust and a number of control variables. Their hypothesis H7 suggests that "perceived risk in the community of sellers decreases intentions to transact with sellers in an online marketplace" (pg. 46). While the structural equation model admits a covariance-equivalent model that reverses the regression relationship of perceived risk on transactional intentions, it is difficult to argue for a plausible corresponding causal relationship in this case. While perceptions can affect intentions, it appears less plausible that intentions affect perceptions.

The study by Malhotra, Kim, and Agarwal (2004) examines antecedents and consequents of internet user's information privacy concerns. Their model includes the constructs Trusting Beliefs and Risk Beliefs, both of which depend on Privacy Concerns and Type of Information, i.e. they have the same antecedents. Malhotra et al. (2004) argue in their H3 that "the more trust a consumer has in an online firm, the less likely he or she is to foresee the risk in providing personal information" (pg. 342). An alternative argument may be that if the perceived risk of loss or damage is low, the trusting belief that "a firm is dependable in protecting consumers' personal information" (pg. 341) might be higher: If I have little of value to lose, I am willing to put more trust in a web-based company. The resulting model, in which the regression relationship between Trusting Beliefs and Risk Beliefs is reversed, is covariance-equivalent to the published model.

In a study on decision aid usage by Hess, Fuller, and Matthew (2005), the authors conclude that the number of decision aid features used has a strong positive effect on the decision time. However, both variables have a single common antecedent, the degree of user involvement in the decision making. As the paper does not offer any theoretical reason for including the regression path in the theoretical model (it is not a formally proposed, theory-based hypothesis), two alternative causal explanations come to mind. First, using decision aids may increase the amount of time spent on decision making. This corresponds to the directionality modelled by Hess et al. (2005). Alternatively, less time spent making a decision may preclude the use of some of the available decision aids. This plausible causal structure corresponds to a reversed regression path in a covariance-equivalent model. While this study was experimental research, both constructs in question were concurrently measured outcome variables, rather than manipulated variables.

A study on project performance by Jiang, Klein, and Chen (2006) contains a model with a single large block within which equivalent models are possible (Figure 7). Jiang et al. (2006) argue in their hypothesis H3 that partnering users and developers within a project team reduces non-support of a project by users. However, the argument that user non-support of a project is the reason that users are partnered with developers in the project also appears to be plausible. In other words, the higher the user non-support, the more inclined project managers will be to involve users and partner them with developers. Similarly, in contrast to the argument that Jiang et al. (2006) make in support of their hypothesis H4, it appears plausible that a high performance risk (difficulty in estimating remaining work) may lead project managers to recognize the benefits of user involvement and partnering. Their model is a fully saturated structural model, in which the $\chi 2$ test does not test any structural constraints and differences to perfect model fit are due only to measurement model constraints. This can be easily seen by replacing all structural regression paths with (undirected) free covariances, yielding a CFA style equivalent model.

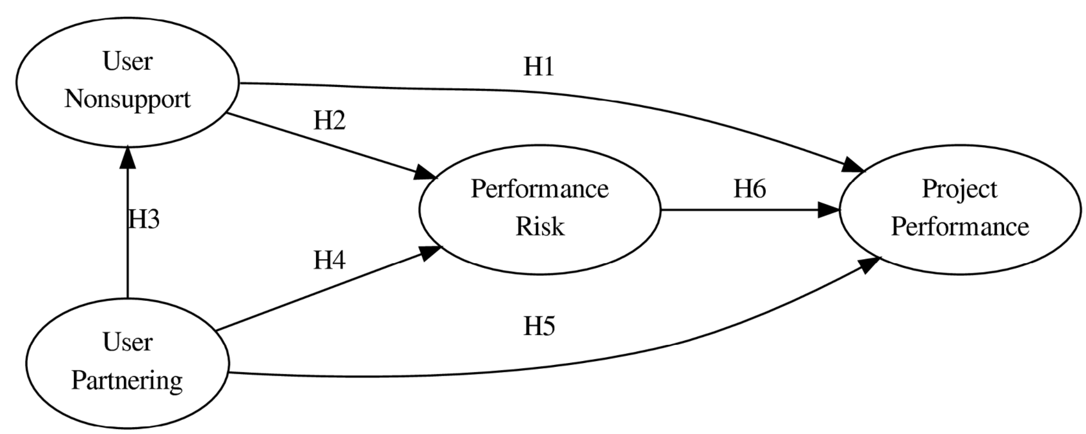

Figure 7: Research Model from Jiang, Klein, and Chen (2006)

A study on willingness to pay for online content, dependent on expected benefits, technical quality and reputation of the vendor and website (Lopes and Galletta, 2006) uses a model with the same structure as that by Jiang, Klein, and Chen (2006) (Figure 7). Consequently, their model also exhibits multiple covariance-equivalent models. For example, Lopes and Galletta (2006) suggest that reputation of an online vendor affects the perceived overall technical quality of the vendor's web site, citing a number of studies in the consumer behaviour and marketing literature. However, it also appears to be plausible that a high-quality web site might improve a vendor's reputation, especially if that vendor is unknown to subjects in the study and subjects have no preconceived notion of the vendor's reputation. This structural model is also fully saturated, in the same way as that by Jiang et al. (2006), thus affording no test of any structural relationship.

A study on manufacturing flexibility tests a model of 5 variables, with a block of four variables being fully connected (saturated), including the constructs Manufacturing Flexibility, Virtual Integration, Supplier Responsiveness, and Environmental Uncertainty (Wang, Tai, and Wei; 2006). Wang et al. (2006) argue in their H2 that “the greater the environmental uncertainty faced by a manufacturer, the greater will be the extent to which it is virtually integrated with its suppliers” (pg. 46). At the same time, they appear to implicitly acknowledge that “virtual integration can be seen as a strategy to reduce the influences of environmental uncertainty". In contrast to their H2, this plausible alternative causal relationship suggests that environmental uncertainty is affected by the level of integration: Higher integration will reduce environmental uncertainty. A similar argument can be made for their H3, where a covariance-equivalent model represents the hypothesis that environmental uncertainty can be affected by a firm's supplier responsiveness.

The extent of internet use in search is argued to affect the extent of internet use in order initiation and completion in a firm's procurement process (Mishra, Konana, and Barua; 2007). However, both constructs share the same set of five exogenous antecedent variables, casting doubt on the directionality of the causal structure. If one assumes that the five exogenous variables also have free covariances, which is a common practice that is often not explicitly stated, the entire block of seven variables is saturated and no specific directionality of the relationships between any two variables is implied by the observed covariances.

The model of IT dependent strategic agility presented by Fink and Neumann (2007) contains a saturated block of four variables, including IT-dependent strategic agility as the dependent variable, IT-dependent system agility, IT-dependent information agility, and infrastructure capabilities. While this block admits multiple covariance-equivalent models, it is more difficult to suggest plausible alternative models than in other cases because the underlying process is less ambiguous: infrastructure typically affects systems, not vice versa; systems affect information, not vice versa, etc. Consequently, a variance theory involving properties of infrastructure, systems, and information built on this process admits little ambiguity.

A study on techno-stress, role-stress, and productivity by Tarafdar, Tu, Ragu-Nathan, and Ragu-Nathan (2007) is in an example where a simple block of three variables may be used to derive equivalent models. The authors first use three separate models to investigate the effects of techno-stress on productivity, techno-stress on role-stress and role-stress on productivity. These smaller models are also ambiguous in representing the causal nature as they only contain two variables, which may be in any relationship with each other. When combining these models, the three variables form a saturated block, allowing no conclusions about the direction of causality from the cross-sectional sample.

Zahedi and Song (2008) describe a study on trust and satisfaction in the healthcare sector. However, while their model contains a saturated block of three variables (Trust Attitude, Extent of Information Use, and Satisfaction), theirs is a longitudinal study where Trust Attitude was measured prior to measuring Extent of Information Use and Satisfaction. This greatly reduces the number of plausible alternative models, leaving only the hypothesized effect of Extent of Information Use on Satisfaction (their H7) in doubt, as these two variables are measured concurrently. Implicitly, Zahedi and Song (2008) appear to acknowledge a possible reversed causality, when they write that "it is satisfaction that builds loyalty by bringing consumers back to the web site" (pg. 232), i.e. satisfaction increases information use because satisfied customers are loyal to the web site.

A study by Kim (2009) on an integrative framework of technology use also admits a covariance-equivalent model, despite being a longitudinal study. A single construct measured at T=1 is the common antecedent of two constructs at T=2 (Technology usage, and TPB determinants). As these are measured concurrently, the directionality of the causal relationship between them cannot be established using the model by Kim (2009).

Most recently, a study by Beaudry and Pinsonneault (2010) presents a model of IT use with antecedent variables Anger, Anxiety, Venting, Seeking Social Support, and Distancing. Venting, Seeking Social Support and Distancing have the same antecedents, Anger and Anxiety. Thus, the model contains two blocks within which the modelled causality is questionable. Specifically, this casts doubt on the conclusions by Beaudry and Pinsonneault (2010) that both Venting and Distancing affect Seeking Social Support. Beaudry and Pinsonneault (2010) are apparently aware of the reverse alternative, pointing this out in their footnote 5: "this type of social support usually amplifies feelings of anger" (pg. 695).

In addition to these studies, we identified another seven studies that are potentially subject to covariance-equivalent model problems. The models in these studies are structured in such a way that a set of two or more exogenous variables determine a single consequent variable (which may have other consequents of its own). If the covariances between exogenous latent variables are free, these models admit covariance-equivalent solutions with reversed direction of regression paths (similar to the top-most model in

Figure 5). This is a common assumption in covariance-based structural equation modeling, but was not explicitly stated in these seven studies.

Finally, we found numerous studies that admit “trivial” covariance-equivalent models, not included in the count of 22 studies we mentioned previously. In these models, two or more exogenous variables affect a single consequent variable. In contrast to the just described seven studies, this consequent also has other antecedent variables. In these cases, only the covariances between exogenous latent variables can be replaced by directional regression relationships, representing directional causal effects. This is well-known to most researchers. We term this “trivial” because these relationships among exogenous latent variables are not usually of theoretical interest.

We acknowledge that the plausibility of the reverse effects we have discussed for these studies is subjective and that some readers may find the arguments in the original studies to be more compelling than our alternatives. However, we emphasize that these issues cannot be settled by argument but are empirical questions and must be settled empirically. Our point here is that the statistical models that are used in these studies cannot do this.

Conclusions and Recommendations

In summary, the detailed example and brief discussion of alternative models for studies in prominent IS journals show that equivalent models can substantially affect the validity of conclusions that can be drawn from a covariance analysis based study. These examples illustrate the difficulty of identifying the causal nature using cross-sectional survey studies, unless care is taken when specifying theories and the consequent statistical models.

When a study cannot rule out alternative explanations, as is the case with studies for which equivalent models can be constructed, its internal validity is threatened (Straub, Boudreau, and Gefen, 2000). While we have not critiqued the theoretical reasoning of every study detail, it is sufficient to present plausible alternative causal relationships to cast doubt on the validity of conclusions drawn from them. As is evident from the previous section, a significant number of theoretical conclusions in IS research are threatened by the existence of covariance-equivalent models. We acknowledge that it may sometimes be the case that covariance-equivalent models, and the threat to the validity of conclusions they pose, is recognized by the researchers but not discussed due to page limitations by journals. However, as the threats frequently go the core of the presented theory, we believe the issue is sufficiently important to deserve mention, if not discussion and resolution.

To address the problem, our first recommendation must be that IS researchers become aware of potential problems. To this end, researchers should examine their models, identify saturated blocks of variables with identical antecedent variables and apply the rules by Lee & Hershberger (1990) to generate covariance-equivalent models. Even when a given model admits multiple equivalent models, this does not mean that all are equally plausible. For example, we noted the studies by Li et al. (2005) and Fink and Neumann (2007) that included constructs for which a strong case could be made for only one direction of causality. In Li et al.'s (2005) study, it was implausible that innate personal characteristics are affected by situation specific beliefs about objects, whereas in Fink and Neumann's (2007) study, it was implausible that IT infrastructure agility is affected e.g. by IT strategic agility. Including such variables in the study can reduce the risk posed by equivalent models.

A second recommendation is to increase the use of longitudinal data or experimentally manipulated variables. For example, we noted the studies by Zahedi and Song (2008) and Kim (2009) whose models admit covariance-equivalent models. The number of plausible equivalent models is greatly reduced by their use of longitudinal data. Longitudinal data and experimental manipulation of variables allow researchers to make a much stronger case for the hypothesized causality of their theoretical relationships. However, as our discussion of the experimental study by Hess et al. (2005) showed, multiple concurrently measured outcome variables may still admit plausible alternative equivalent models.

A third recommendation is to structure the statistical model in such a way to preclude covariance-equivalent models in the first place. We do not mean to suggest that statistical considerations should drive the research. While theoretical considerations must drive statistical choices, researchers need to be aware of the limitations of the statistical techniques in testing theoretical models. In the case of covariance-based analysis, these limitations may preclude researchers from testing certain aspects of their theories, for which covariance-equivalent alternative models exist. Thus, we do not ask researchers to look to statistics over theory, but to be aware of the limitations and constraints of the chosen statistical approach. In the more general sense, we need to ensure that our theories and their hypotheses remain empirically testable or falsifiable, which is an important, if not essential quality of scientific theories (Bacharach, 1989). Statistics sets limits on the testability of hypotheses and theories, and we must ensure that the theories we develop stay within these limits. Both the complexities of social contexts, and the limits of the statistical tools available make it likely that IS researchers will develop theories that have plausible covariance-equivalent models. A major contribution of this paper is to enable researchers to recognize when (aspects of) the hypotheses represented in these models are not testable using covariance analysis. It is then incumbent upon the researcher to use longitudinal or experimental studies to examine the problematic hypotheses.

One way of appropriately structuring the statistical model is through the use of additional control variables, i.e. variables of no theoretical importance, might be introduced to the model in such a way as to reduce the number of covariance-equivalent models. For example, Pavlou and Gefen (2004) might have precluded a covariance-equivalent model simply by adding Trust Propensity as a control variable also to Perceived Risk, not only to Trust.

Another way of structuring the statistical model is to recognize that the common practice in covariance-based studies to free the covariances among exogenous latent variables for estimation contributes to the creation of equivalent models. For an example, consider Figure 5. In the top-most model in Figure 5, variables A and C are common antecedents to B and are related by a free covariance. Hence, we can generate 12 covariance-equivalent models for this set of three variables. If however, the model assumed that A and C are independent, i.e. fixed the covariance at 0, the model would not admit equivalent models. Researchers should look critically at the independence of the exogenous latent variables in the model and impose constraints if possible.

Fourth, as we noted in our brief section on covariance-equivalence, saturated models are more likely to admit plausible covariance-equivalent models than sparser models. We identified a number of studies with a completely saturated structural model (Jiang et al., 2006; Lopes and Galletta, 2006; Tarafdar et al., 2007). These models are equivalent to CFA models and do not test any of the structural regressions, as these can be arbitrarily replaced by other relationship types. Theoretically, it is easy to free a model parameter for estimation (i.e. adding a path) as the statistical significance of the path parameter estimate can be tested anyway. Adding paths also typically improves the model fit. However, the price to be paid is the possible existence of covariance-equivalent models and the consequent lack of validity of the theoretical conclusions. We recommend that researchers use strong theory if at all possible to fix path coefficients based on prior theory or remove paths altogether when their existence is not strongly dictated by theory. It is useful to remember that the $\chi 2$ model test is a test of the model constraints (e.g. the lack of paths), not the model paths: The $\chi 2$ statistic is used to test how well a model fits the observed covariance information given the constraints imposed on it. We recommend an increased emphasis on theory testing that focuses on whether the imposed constraints hold (Evermann and Tate, 2011), instead of focusing on the statistical significance of freely estimated parameters.

Finally, the issue of equivalent models has been raised specifically in covariance analysis because this technique provides an overall test of model fit, which, while very useful, cannot differentiate among covariance-equivalent models. However, this does not imply that other techniques, such as regression or partial least squares analysis, are immune from these issues. For a more general discussion of the relationship between statistics and causality, the reader is referred to the thorough treatment by Pearl (2009).

In regression analysis, the issue of "reverse causality" has been addressed in the econometrics literature, where regression analysis is dominant, using the technique of instrumental variables (Angrist and Krueger, 2001). Instrumental variables address the issue of endogeneity, where the predictor variable is correlated not only with the dependent, but also with the regression error term. Such endogeneity is an indicator that the regression is ambiguous with respect to the direction of causality. An instrumental variable is one that is correlated with the dependent variable, but not the error term. Instrumental variable regression models can be estimated using the 2-stage least squares (2SLS) estimation method.

While 2SLS can also be used to estimate SEM models (Kirby and Bollen, 2001), in the SEM context, the instrumental variable can be explicitly and easily included in the model (Antonakis et al., 2010).

In the context of Partial Least Squares analysis, common in IS research, there has been no discussion of the issue of reverse causality. However, given the lack of strict statistical tests of overall model fit, the potential problems for these techniques are harder to identify than for covariance-based statistical models. Because of the lack of a test of model fit (Evermann and Tate, 2010), correctness of models (and therefore discrimination between equivalent models) is impossible to determine. For example, the PLS model for the detailed example by Luo et al. (2005) shows a goodness-of-fit (Chin, 2010) of 0.9683 whereas the PLS fit for the covariance-equivalent model is 0.9732, a minor difference that leaves both values above the recommended 0.9, demonstrating that PLS is not immune to the problem of statistically equivalent models.

Table 2 shows our specific recommendations and step-by-step guidance to recognize and address the issue of covariance equivalent models.

Step	Recommendation
1	Construct theoretical model based on substantial considerations
2	Construct statistical model from theoretical model
3	Examine statistical model for saturated blocks that indicate the possibility of covariance-equivalent models within those blocks
For each saturated block	
For each path within the block	
Either	
4a	Identify prior experimental or longitudinal literature to support a specific direction.
4b	Add additional predictor variable(s) to one of the path nodes to avoid covariance equivalence. These variables do not need to have theoretical importance for the tested theory but should have theoretical rationale.
4c	Remove path from statistical model → the corresponding theoretical hypothesis cannot be statistically tested.

Table 2: Recommendations to identify and avoid covariance-equivalent models

In conclusion, we believe the problem of covariance-equivalent models is a serious issue that affects the validity of theoretical conclusions and practical recommendations of a significant body of work in the IS literature. We encourage IS researchers to attend to this issue through problem identification, appropriate model construction, and longitudinal and experimental studies.

References

Angrist, J. and Krueger, A. 2001. "Instrumental variables and the search for identification: From supply and demand to natural experiments." *Journal of Economic Perspectives* (15:4), 69–85.

Antonakis, J., Bendahan, S., Jacquart, P. and Lalive, R. 2010 "On making causal claims: A review and recommendations." *The Leadership Quarterly* (21:6), 1086-1120.

Bacharach, S.B. 1989. "Organizational theories: Some criteria for evaluation." *The Academy of Management Review* (14:4), 496-515.

Beaudry, A. and Pinsonneault, A. 2010. "The other side of acceptance: Studying the direct and indirect effects of emotions on information technology use." *MIS Quarterly* (34:4), 689-710.

Chin, W.W. 1998. "Issues and opinion on structural equation modeling." *MIS Quarterly*(22:1), vii-xvi.

Chin, W.W., Johnson, N., and Schwarz, A. 2008. "A fast form approach to measuring technology acceptance and other constructs." *MIS Quarterly* (32:4), 687-703.

Chin, W.W. 2010. "How to write up and report PLS analyses". In: Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (eds.) *Handbook of Partial Least Squares*. New York: Springer Verlag.

Evermann, J. and Tate, M. 2010. "Testing Models or Fitting Models? Identifying Model Misspecification in PLS". *Proceedings of the 31st International Conference on Information Systems ICIS, St. Louis, MS.*

Evermann, J. and Tate, M. 2011. "Fitting covariance models for theory generation." *Journal of the Association for Information Systems* (12:9), 632-661.

Fink, L. and Neumann, S. 2007. "Gaining agility through IT personnel capabilities: The mediating role of IT infrastructure capabilities." *Journal of the AIS* (8:8), 440-462.

Hayduk, L.A. 1996. *LISREL Issues, Debates, and Strategies*. Baltimore, ML: Johns Hopkins University Press.

Hess, T.J., Fuller, M.A., and Mathew, J. 2005. "Involvement and decision-making performance with a decision aid: The influence of social multimedia, gender, and playfulness." *Journal of Management Information Systems* (22:3), 15-54.

Jiang, J.J., Klein, G., and Chen, H.-G. 2006. "The effects of user partnering and user non-support on project performance." *Journal of the AIS* (7:2), 68-90.

Kim, S.S. 2009. "The integrative framework of technology use: An extension and test." *MIS Quarterly* (33:3), 513-537.

Kirby, J.B. and Bollen, M.A. 2001. "Using instrumental variable (IV) tests to evaluate model specification in latent variable structural equation models." *Sociological Methodology* (39:1), 327-355.

Lee, S., and Hershberger, S. 1990. "A simple rule for generating equivalent models in covariance structure modeling." *Multivariate Behavioral Research* (25:3), 313-334.

Li, D., Chau, P.Y.K., and Lou, H. 2005. "Understanding individual adoption of instant messaging: An empirical investigation." *Journal of the AIS* (6:4), 102-129.

Maccallum, R.C., Wegener, D.T., Uchino, B.N., and Fabrigar , L.R. 1993. "The problem of equivalent models in applications of covariance structure analysis." *Psychological Bulletin* (114:1), 185-199.

Malhotra, N., Kim, S.S., and Agarwal, J. 2004. "Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model." *Information Systems Research* (15:4), 336-355.

Markus, K. 2002. "Statistical equivalence, semantic equivalence, eliminative induction, and the Raykov-Marcoulides proof of infinite equivalence." *Structural equation modeling* (9:4), 503-522.

Mishra, A.N., Konana, P., and Barua, A. 2007. "Antecedents and consequences of internet use in procurement: An empirical investigation of U.S. manufacturing firms." *Information Systems Research* (18:1), 103-120.

Mulaik, S.A. 2009. *Linear Causal Modeling with Structural Equations*. Boca Raton, FL: Taylor & Francis Group.

Parboteeah, D.V., Valacich, J.S., and Wells, J.D. 2009. "The influence of website characteristics on a consumer's urge to buy impulsively." *Information Systems Research* (20:1), 60-78.

Pavlou, P.A. and Gefen, D. 2004. "Building effective online marketplaces with institution-based trust." *Information Systems Research* (15:1), 37-59.

Pearl, J. 2009. *Causality: Models, Reasoning, and Inferences*. 2nd edition. Cambridge, MA: Cambridge University Press.

Raykov, T., and Marcoulides, G. 2001. "Can there be infinitely many models equivalent to a given covariance model?" *Structural Equation Modeling* (8:2), 142–149.

Raykov, T., and Marcoulides, G. 2007. "Equivalent structural equation models: A challenge and responsibility." *Structural Equation Modeling* (14:4), 695-700.

Stelzl, I. 1986. "Changing a causal hypothesis without changing the fit: Some rules for generating equivalent paths models." *Multivariate Behavioral Research* (21:3), 309-331.

Straub, D., Boudreau, M.-C., and Gefen, D. 2004. "Validation guidelines for IS positivist research." *Communications of the AIS* (13), article 24.

Tarafdar, M., TU, Q., Ragu-Nathan, B.S., and Ragu-Nathan, T.S. 2007. "The impact of technostress on role stress and productivity." *Journal of Management Information Systems* (24:1), 301-328.

Venkatesh, V. 2000 "Determinants of perceived ease of use: Integrating perceived behavioral control, computer anxiety and enjoyment into the technology acceptance model." *Information Systems Research* (11:4), 342-365.

Wang, E.T.G., Tai, J.C.F., and Wei, H.-L. 2006. "A virtual integration theory of improved supply-chain performance." *Journal of Management Information Systems* (23:2), 41-64.

Zahedi, F., and Song, J. 2008. "Dynamics of trust revision: Using health infomediaries." *Journal of Management Information Systems* (24:4), 225-248.